CHILD DEVELOPMENT

Child Development, May/June 2014, Volume 85, Number 3, Pages 1062–1076

Deconstructing Building Blocks: Preschoolers' Spatial Assembly Performance Relates to Early Mathematical Skills

Brian N. Verdine and Roberta M. Golinkoff
University of Delaware

Kathryn Hirsh-Pasek and Nora S. Newcombe
Temple University

Andrew T. Filipowicz and Alicia Chang
University of Delaware

This study focuses on three main goals: First, 3-year-olds' spatial assembly skills are probed using interlocking block constructions (N = 102). A detailed scoring scheme provides insight into early spatial processing and offers information beyond a basic accuracy score. Second, the relation of spatial assembly to early mathematical skills was evaluated. Spatial skill independently predicted a significant amount of the variability in concurrent mathematical performance. Finally, the relation between spatial assembly skill and socioeconomic status (SES), gender, and parent-reported spatial language was examined. While children's performance did not differ by gender, lower SES children were already lagging behind higher SES children in block assembly. Furthermore, lower SES parents reported using significantly fewer spatial words with their children.

The child amidst his baubles is learning the action of light, motion, gravity, muscular force.

—Ralph Waldo Emerson

Spatial thinking, or mentally manipulating information about the structure of the shapes and spaces in one's environment, may be crucial for developing skills that support later science, technology, engineering, and mathematics (STEM) learning (Newcombe, 2010; Wolfgang, Stannard, & Jones, 2001). Research shows that spatial skills are malleable (Uttal et al., 2013) and that early experiences like block building in 4- to 6-year-olds (Casey et al., 2008) and puzzle play in 2- to 4-year-olds (Levine, Ratliff, Huttenlocher, & Cannon, 2012) can alter spatial thinking. Complex mathematical problem solving also rests on spatial skills and early links between spatial and mathematical skills have been established (Gunderson, Ramirez, Beilock, & Levine, 2012). Because spatial and mathematical skills are vitally important for success in STEM careers and because spatial skills are foundational to mathematics learning, spatial instruction has become a priority in early education (National Council of Teachers of Mathematics [NCTM], 2007). Block building, in particular, offers a potential route to study and improve these skills in children prior to formal mathematics instruction.

Block Play as a Window Into Spatial Understanding

There are two main types of play with blocks. In *free play* (e.g., as used in Stiles-Davis, 1988), children are given blocks and build designs of their choice. In *structured block play* (e.g., Stiles & Stern, 2009), children attempt to make a particular structure from a model. These tasks tend to call on different processes: The former invokes children's imagination and ability to produce complex relations without prompting, while the latter calls upon the ability to analyze a spatial representation to create

This research was supported by the National Institutes of Health Stimulus Grant 1RC1HD0634970-01 to Roberta Michnick Golinkoff at the University of Delaware and Kathryn Hirsh-Pasek at Temple University, and the National Science Foundation via the Spatial Intelligence and Learning Center (SBE-1041707). The authors thank Marcia Halperin and Nancy Jordan for their consultation on this project; Angeliki Athanasopoulou, ToriAnne Davies, Raissa Dempsey, and Gabrielle Farmer for research assistance and feedback on the manuscript; and Herbert Ginsburg and Sandra Pappas for their help with the EMAS assessment. We also thank Mega Bloks® who provided the blocks used in the study.

Correspondence concerning this article should be addressed to Brian N. Verdine or Roberta Golinkoff, School of Education, University of Delaware, Willard Hall, Newark, DE 19716. Electronic mail may be sent to verdine@udel.edu or roberta@udel.edu.

© 2013 The Authors
Child Development © 2013 Society for Research in Child Development, Inc.
All rights reserved. 0009-3920/2014/8503-0017
DOI: 10.1111/cdev.12165

a predefined model (Stiles & Stern, 2009). Structured block play has been hypothesized to develop skills in estimation, measurement, patterning, part–whole relations, visualization, symmetry, transformation, and balance (Casey & Bobb, 2003). The concept of measurement, for example, is involved when a child compares the height of their construction with the height depicted in a model or when two children argue about how to figure out whose tower is "bigger" (e.g., should they use length, area, and volume, or the number of blocks; Cross, Woods, & Schweingruber, 2009). Thus, structured block play is an excellent venue for studying spatial skills in young children and is the method chosen for this investigation.

Caldera et al. (1999) found an association between block-building skills and spatial visualization. According to Linn and Petersen (1985), spatial visualization involves complicated, multistep manipulations of spatial information and includes multiple solution strategies. Caldera et al. measured it with the block design subtest from the Wechsler Preschool and Primary Scale of Intelligence (WPPSI) and the Wechsler Intelligence Scale for Children–R (WISC–R). They also coded children's behaviors during free block play and found that task engagement and the creation of complex structures during block play were related to success on another spatial task—the embedded figures test (Karp & Konstadt, 1971). Casey et al. (2008) also found that children's spatial skills on the block design subtest of the WISC–IV benefited from goal-oriented block-building activities and that a story context (a puppet dragon helping a king and queen rebuild parts of a castle) was associated with further improvement. The story provided children with reasons for *why* the castle walls should not have gaps or be a certain height, and provided a motivation for adhering to the parameters. It is also possible that the story exposed children to more spatial language (e.g., above, below, on top of, next to, etc.).

Uttal et al. (2013), building on Chatterjee (2008), discuss a typology of spatial skills that provides a framework for relating the outcomes of a structured block task to the existing spatial literature. Their framework is based on two pairs of dimensions that, when crossed, create a 2 × 2 grid (intrinsic vs. extrinsic and static vs. dynamic). Spatial activities that involve distinguishing the characteristics of a single object (e.g., is it big or little?) tap the *intrinsic* dimension, while determining the relation between objects (e.g., is an object next to or on top of another object?) taps the *extrinsic* dimension. When the main object in a spatial array remains stationary (as in a key on a table) it is *static*; when it is moved, either physically or mentally (to open a door), it is *dynamic*. Block building calls upon all of these dimensions. Builders must distinguish between the different, intrinsic properties of static blocks and place them dynamically into new extrinsic spatial arrangements.

Children's Early Mathematical Skills

Since Piaget, researchers have been interested in children's counting and construal of number words. While Piaget and Inhelder (1969) emphasized what children in the preoperational stage could not do, Gallistel and Gelman (1992) highlighted the fact that even preschoolers seem to function with a set of principles for counting. Understanding that numbers are ordinally organized and not just nominal indicators takes time, and children's mathematical experiences undoubtedly influence how quickly they come to appreciate such principles. For example, Saxe (1979) reported that even though children may be able to count, at first they do not use the outcome of their counting procedure to solve number conservation tasks—acting as if numbers are names and not identifiers for quantity. Children from lower socioeconomic environments appear to particularly lag in their understanding of number words (NCTM, 2007).

Spatial Skills and Mathematics

Spatial skills are not only important in their own right but also because they are related to mathematical achievement (Clements & Sarama, 2007; Robinson, Abbott, Berninger, & Busse, 1996) and likely play a *causal* role in improving mathematical skills (Newcombe, 2010; Mix, Moore, & Holcomb, 2011). A recent study by Grissmer et al. (2013) provided experience to kindergarten and first-grade children with sets of visuospatial toys (e.g., Legos®, Wikki Stix®, pattern blocks, etc.) that required them to copy model designs. Although these activities were not specifically mathematical in nature, they improved children's mathematical skills. Another study by Cheng and Mix (2012) showed that training on a mental rotation task with 6- to 8-year-olds improved performance on calculation problems and the use of place value concepts. Kamii, Miyakawa, and Kato (2004) also found that a range of skills that are both spatial and mathematical in nature appear to develop concurrently in children's block building. Furthermore, Gobel, Walsh, and Rushworth (2001) found that similar areas in the brain respond when individuals engage in space and mathematics processing (see Umiltà, Priftis, & Zorzi, 2009, for a review). Thus, visual-spatial skills and mathematics, traditionally taught separately, regularly call on a shared set of foundational skills and may have a significant amount of overlap.

A complete understanding of the relation between learning of mathematics and spatial skills requires further investigation, but we do have clues about why these domains are related. First, preverbal children use number magnitudes and subitizing strategies to "count" an array. As children begin to learn words and count, this system requires refinement to create the one-to-one mapping between objects and the discrete magnitudes of number words (Gallistel & Gelman, 1992). Second, Gunderson et al. (2012) reported that spatial skills in kindergarten and early elementary school play a role in how well children perform on a linear number line task between 6 months and 1 year later. Understanding that the distance between numbers in a spatial representation maps on to the magnitude of those numbers is a fundamentally spatial task. Counting and using discrete magnitudes is something inherent in blocks, especially popular blocks like Legos® that have pips that attach the pieces together. Placing these pieces correctly requires counting or an understanding of units and invokes measurement concepts. Practice with blocks may provide an early analogue for learning explicit measurement concepts and for understanding discrete units, helping build a more concrete link between number magnitudes and number language.

There are some indications that the link between early and later spatial and mathematical skills becomes stronger as children get older (e.g., Battista, 1990). This is something of a paradox since typically correlations between related skill areas are higher when assessed closer in time. A possible explanation is that *earlier* mathematics tests tend to require problem solving based on memory (e.g., $4 + 3 = 7$) and clear-cut algorithms like "order of operations" (e.g., $4 + 3 \times 6 = 4 + 18 = 22$), rather than generation and interpretation of spatial representations (Wolfgang et al., 2001). Therefore, skills like mental rotation and spatial visualization may not be invoked. It may also be the case that those with good spatial skills rise to the top in mathematics as the curriculum becomes more complicated and incorporates more spatial elements. One study found strong correlational links between space and mathematical skill in a large sample of mathematically precocious young students (Robinson et al., 1996). Research with more typical samples—as tested here—is necessary to understand the early relation between spatial and mathematical thinking. We predicted moderate correlations between spatial and mathematical skills at 3 years of age.

Spatial Skill, Socioeconomic Status (SES), and Gender

Previous research indicates that higher SES children outperform lower SES children on figure drawing and modeling (Fuson & Murray, 1978). Such differences are likely a result of variable experiences at home and in school (e.g., Burchinal, Nelson, Carlson, & Brooks-Gunn, 2008; Hart & Risley, 1995, 2003; Jordan, Kaplan, Nabors Oláh, & Locuniak, 2006). Although some school curricula, including those for Head Start, specifically incorporate some mathematics and block-building activities, there is a significant amount of variability in the instruction.

Some researchers have argued that gender differences in spatial skill are also the result of dissimilar experiences and that these differences grow larger in older children (Voyer, Voyer, & Bryden, 1995) and vary by sample SES (Levine, Vasilyeva, Lourenco, Newcombe, & Huttenlocher, 2005). On the other hand, gender differences in specific aspects of spatial skill are consistently found (e.g., mental rotation; Linn & Petersen, 1985), even as early as 3–4 months of age (Moore & Johnson, 2008; Quinn & Liben, 2008).

One experience that may be different for SES and gender groups is their exposure to spatial language. Block play elicits more spatial language than nonspatial play (Ferrara, Hirsh-Pasek, Newcombe, Golinkoff, & Shallcross Lam, 2011) and spatial language appears to influence the development of spatial cognition. For example, in a longitudinal study by Pruden, Levine, and Huttenlocher (2011), children hearing more spatial language (words such as big, tall, and circle) from 14 to 46 months performed better on spatial tests at 54 months. In Levine et al. (2012), six observations of parent–child interactions during typical daily activities were made from 26 to 46 months followed by a test of spatial transformation skill at 54 months. The frequency of puzzle play—an elicitor of parental spatial language—was associated with better spatial task performance in females.

The Present Study

To probe the relation between spatial and mathematical skill in preschool children, this study asks three questions not addressed in prior research. First, can we develop an age-appropriate task for

3-year-olds that tests their spatial assembly skills? Second, does spatial assembly relate to emerging mathematical skills even as young as age 3? Finally, are SES and gender associated with differential spatial and mathematical skill levels at this early age?

Spatial skills were Assessed via a Test of Spatial Assembly (TOSA) using interlocking, colored plastic blocks called Mega Bloks®, larger than Legos® and more appropriate for young children. Children were asked to build a set of target constructions from models (see Table 1) and were given identical pieces to the model structure. This test was created because there are few spatial tests for this age group and those that do exist may not provide sufficient insight into children's spatial thinking. Standardized tests like the WPPSI block design and the Woodcock–Johnson III Spatial Relations subtest (Woodcock, Mather, & McGrew, 2001) focus on the products of learning rather than the *processes* children use (Hirsh-Pasek, Kochanoff, Newcombe, & de Villiers, 2005). Two children with similar scores on these tests may have made different errors and the scoring rubrics are blind to such variations. Furthermore, the WPPSI uses cubes of the same size, preventing an understanding of how children deal with the metric properties of blocks of varying lengths. The prior research that has focused on *how* children assemble blocks (Casey et al., 2008; Stiles & Stern, 2009; Stiles-Davis, 1988) did not relate these processes to concurrent or later measures of spatial and mathematical skills in children this young. The TOSA overcomes these problems.

In copying the three-dimensional (3-D) block designs from the TOSA (see Table 1), children must first observe which level (of two levels) a block is placed on. Our detailed coding scheme captured this as the *vertical location* dimension. Second, children must note how each block is oriented relative to the block it sits on (*rotation* dimension). That is, does the top block share the same orientation as the bottom block or is it oriented perpendicular to it? Finally, the precise location of a block relative to the position of other blocks must be gauged (*translation* dimension).

To investigate the relation between spatial assembly and mathematical skill, we also gave the Early Mathematics Assessment System (EMAS; Ginsburg, Pappas, & Lee, 2012) and collected parent-reported data about the spatial relational terms parents used with their child. Parent reports of spatial language are expected to predict spatial assembly scores. The Peabody Picture Vocabulary Test (PPVT) was used to control for the influence of verbal ability on the relations between space and mathematical skills.

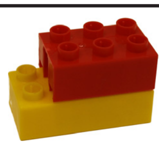
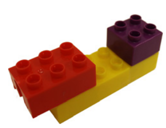
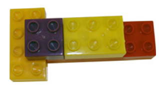
Finally, information about the educational level of the mother (a proxy for SES) and the child's gender allowed us to assess their relation to spatial and mathematics performance. Considering previous research, we expected that lower SES children would score below their higher SES peers, but that gender differences might not emerge.

Method

Participants

Participants were 102 preschoolers (55 boys, 47 girls) between the ages of 38 and 48 months (M = 43.6 months) recruited from Head Start facilities (N = 42; with all but 1 child classified as lower SES), private preschools (N = 31; all but 1 child classified as higher SES), and a university-related (UR) preschool (N = 26; 18 higher SES and 8 lower SES). All

Table 1
Dimensional Scoring: Points Possible on the 3-D TOSA for Each Design by Dimension

Design	 1	 2	 3	 4	 5	 6	Total
No. of pieces	2	2	3	3	4	4	18
Block lengths	2, 3	3, 4	2, 3, 6	2, 3, 4	2, 3, 4, 6	2, 3, 4, 6	
Vertical location	1	1	3	3	5	6	19
Rotation	0	1	1	1	2	3	8
Translation	1	1	3	3	5	6	19
Total	2	3	7	7	12	15	46

Note. Piece rotation was not scored for component blocks that are 2 pips × 2 pips because they are symmetrical. TOSA = Test of Spatial Assembly; 3-D = three dimensional.

children were native English speakers and had no identifiable vision or language problems.

Parents returned a 14-item survey that requested mother's educational level, a variable that Hoff (2013) argued may be the most critical SES component for development. Creating two SES groups based on mothers' education had the side effect of sorting all but two of the children by type of preschool attended (the UR preschool was mixed, containing children of both faculty and staff). Mothers of the higher SES group (24 boys, 27 girls) completed a bachelor's or graduate degree; the lower SES group (31 boys, 20 girls) had some high school education up to an associate's degree.

The lower SES group was slightly older than the higher SES group on average (M = 44.2 vs. 43.0), $t(100) = 2.29$, $p = .024$, but age at testing was not correlated with performance on any of the measures. Where parental survey data were not returned, children from Head Start centers were categorized in the lower SES group ($N = 3$) and children from private preschools in the higher SES group ($N = 6$).

Stimuli

The 3-D TOSA included seven constructions (one training, six test trials) made of interlocking Mega Bloks® modified by cutting length off of a block or gluing blocks together to provide more size options. Constructions included two to four blocks. Each construction (Table 1), was made of 24 mm high blocks ranging in length from one unit (2 pips × 2 pips; 32 mm × 32 mm; pips are the knobs that lock the pieces together) to three units (2 × 6; 32 mm × 96 mm).

Items were selected to provide a range of difficulty. The number of blocks on the top row compared to the bottom was varied. The orientation of blocks in relation to the base piece was also varied (e.g., see Items 3 and 4). Finally, our harder designs incorporated more instances where a block was only attached by half of the width of another block (i.e., by only two pips), requiring even more fine-grained distinctions about relative location.

Procedure

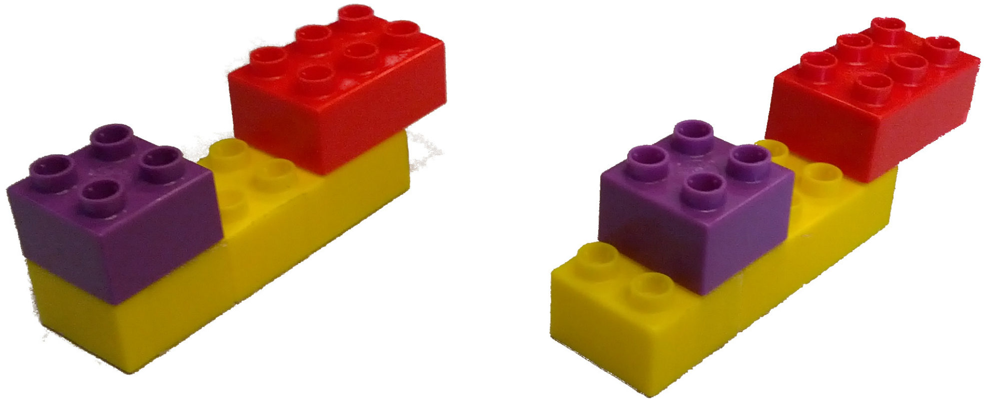
Children were individually tested in a quiet room seated across the table from the experimenter. Children started with a training trial in which two blocks were placed on the table and the experimenter told children that they were going to make "these pieces look just like this one [the model]." The experimenter then connected the blocks incorrectly, asked the child to confirm whether the pieces matched the model, and pointed out that there was a mistake if the child did not recognize one. The blocks were then rearranged to correctly match the model and the child confirmed the match. Finally, the experimenter disconnected the blocks, placed the pieces in front of the child, and the child was asked to "make the pieces look like this one [pointing to the model]." All participants succeeded on the training trial with their first attempt.

In each test trial, the experimenter presented the child with a 3-D, intact, glued-together model construction and a matching set of separated bricks arranged randomly. The child was then asked, "Can you make your pieces look just like this one?" During the *untimed* task, the experimenter waited for children to indicate that they were finished. The model was always left on the table in plain sight and children were allowed to pick it up, although they were discouraged from playing with it and were not allowed to add their pieces to it. The number and type of blocks the child was given were the same as those used in the model. When the child was finished, the experimenter removed the construction and introduced the next item. Participants received all six test trials in the same order from easier to harder models. Difficulty order was established by the consensus of lab personnel based on the number of blocks, rotated components, overhanging blocks, and whether the designs had multiple blocks of the same color. Pilot testing confirmed the expected order was generally accurate. Photographs were taken of the completed constructions for later coding.

3-D TOSA Scoring

There were two types of scoring: Match scoring and Dimensional scoring (Figure 1). Match scoring gives participants 1 point for each of the six designs for which they *matched* the model *100% correctly*. This scheme provided an easy-to-code metric of *performance* on the task and a point of comparison for our Dimensional scoring.

Dimensional scoring codes specific errors and was devised to uncover the *processes* children use and the mistakes made in recreating block models. The first step of Dimensional scoring rated accuracy relative to a central piece in the design—called the *base*. The base was the biggest piece or the piece that had the most other pieces attached. One point was awarded for each component piece on each of



Model Construction Child's Construction

Figure 1. Test of spatial assembly coding example for dimensional scoring, Steps 1 and 2. Dimensional coding (total points = 5 of 7): *Step 1*: Base-relative coding (3 points)—The location of the two component pieces are coded relative to the yellow, 6 × 2 pip base piece. Vertical location (child awarded 2 points)—Both pieces are one level above the base piece, as in the model. Rotation (awarded 1 point)—2 × 2 blocks, like the purple one to the left in the constructions, are symmetrical and therefore not scored for rotation. The red 2 × 3 block on the upper right of the constructions receives a point because the main axis of the block is oriented parallel with the main axis of the base piece in both constructions. Translation (awarded 0 points)—Both component blocks are shifted one set of pips on the base piece. *Step 2*: Dyad Coding (2 points)—One dyad appears: the red 2 × 3 block and the purple 2 × 2 block. The bigger 2 × 3 block serves as the ground piece. Vertical location (awarded 1 point)—The dyad blocks remain on the same level in the child's construction as in the model. Rotation (awarded 0 points)—The 2 × 2 piece is not scored because it is symmetrical. Translation (awarded 1 point)—The component pieces maintain a 2 pip × 2 pip space between them.

three independently scored dimensions. A point for *vertical location* was awarded if a component block was on the correct level of the design compared to the base. *Rotation* was scored by determining if a piece's axis was oriented correctly with respect to the long axis of the base piece (parallel or perpendicular to it). For example, if the long axes of the component and base pieces were perpendicular in the model and the child copied this orientation, he or she received a point. A *translation* point was awarded if a component piece was placed over the correct pips in relation to the base piece. These dimensions represent the range of errors children could make as they place blocks in 3-D space. That is, *x*- and *y*-axis location errors are captured with the vertical location and translation dimensions; the accuracy of the orientation of the block in its location is captured with the rotation dimension.

For *rotation*, chance performance is 50% as there are only two possible orientations in which the blocks can lock together (perpendicular or parallel to another block). For *vertical location*, the expected performance level for chance responding is inversely related to the number of blocks in the design. For example, a single piece in a three-block design could appear in *up to* five possible places (i.e., one or two levels above, on the same level, and one or two levels below). The *translation* dimension is even less constrained because there are a larger number of locations for a component piece to be placed incorrectly. Therefore, chance responding would yield fewer points. Reliability coding for the first scoring step was done for approximately 20% of the sample (N = 21) with 95.9% agreement.

The second Dimensional scoring step focused on the more complex constructions with multiple component pieces (Designs 3–6) and coded how well those pieces maintained accurate relations with respect to one another. Here, the base piece was ignored; component pieces were coded according to the pairs, or dyads, in which the pieces occurred. Thus, Designs 3 and 4 each contained one dyad, Design 5 contained two dyads, and Design 6, three dyads. One piece from each dyad was designated as the *ground* piece, which was the larger of the two pieces. Dyads were then scored using the same three scoring dimensions described above. This second coding step was included for maintaining some interpiece relations in the more complex designs (3–6). Children could maintain relations between component pieces even if they were not placed properly with respect to the base piece.

The resulting scores from both coding steps were added to create the 3-D TOSA Dimensional scores. Table 1 shows the possible points for each design. Percentage agreement for this coding step (97.2%) was also high.

Additional Measures

Mathematics. The EMAS: Number and Operations subtest (Ginsburg et al., 2012) was given as a test of early mathematical skill. The EMAS included (a) a verbal *free counting* task noted the highest number the child could count to without a mistake; (b) a *Give-N task* (e.g., Wynn, 1990), consisting of four items, asked the child to give a specific number of objects (3, 1, 2, 4) from a larger quantity (7, 4, 6, and 7, respectively); (c) a *number order* task, consisting of three items, indicated sequential knowledge of numerals (e.g., "What number comes after 4 when we count?"; numbers 4, 5, and 7 were used); and (d) a *nonverbal addition and subtraction task* (e.g., Levine, Jordan, & Huttenlocher, 1992), consisting of three items, involved hidden objects that showed mental computation skill. The experimenter would show one chip, put it under a mat, show another chip, and slide it under the mat, requiring addition of the hidden chips. The problems used were 1 + 1, 3 − 1, and 2 + 1. A total of 10 points were possible for the EMAS. The free counting measure, a scale variable, was not incorporated in the total score and was analyzed separately. The reported alpha for this EMAS subtest for 3-year-olds is .86 with a test–retest reliability of .79 (Ginsburg, Lee, Pappas, Hartman, & Rosenfeld, 2010) and path models had a factor loading of .93 on early mathematical skill.

Spatial language exposure. Parents were given a checklist with the words *big*, *between*, *below*, *behind*, *next to*, *short*, *little*, *on*, *above*, *near*, *under*, *in*, *long*, and *in front* and asked, "Which of the following words do you use with this child?" Checking all that applied yielded a dichotomous yes or no response for each word on the list.

Language skill. Participants were given the PPVT—4th ed. (Dunn & Dunn, 2007). Split-half and alpha reliability for 3-year-olds are ≥ 0.93 and alternate form and test–retest correlations are ≥ 0.90. The validity of the PPVT is established through correlations with the Expressive Vocabulary Test–2 (r = .81), Comprehensive Assessment of Spoken Language (r = .41–.54), and Clinical Evaluation of Language Fundamentals–4 (r = .67–.73).

Results

Spatial Assembly on the 3-D TOSA

Match scores. Match scores ask whether the child's design completely matches the model. Similar to the WPPSI scoring system, it does not provide the nuanced analysis Dimensional scoring does. Match scores indicate that this task was suitable for the age group and that the task scaled well. Children had near-ceiling performance on Designs 1 and 2, and 0% correct on Item 6 (see Figure 2). Of

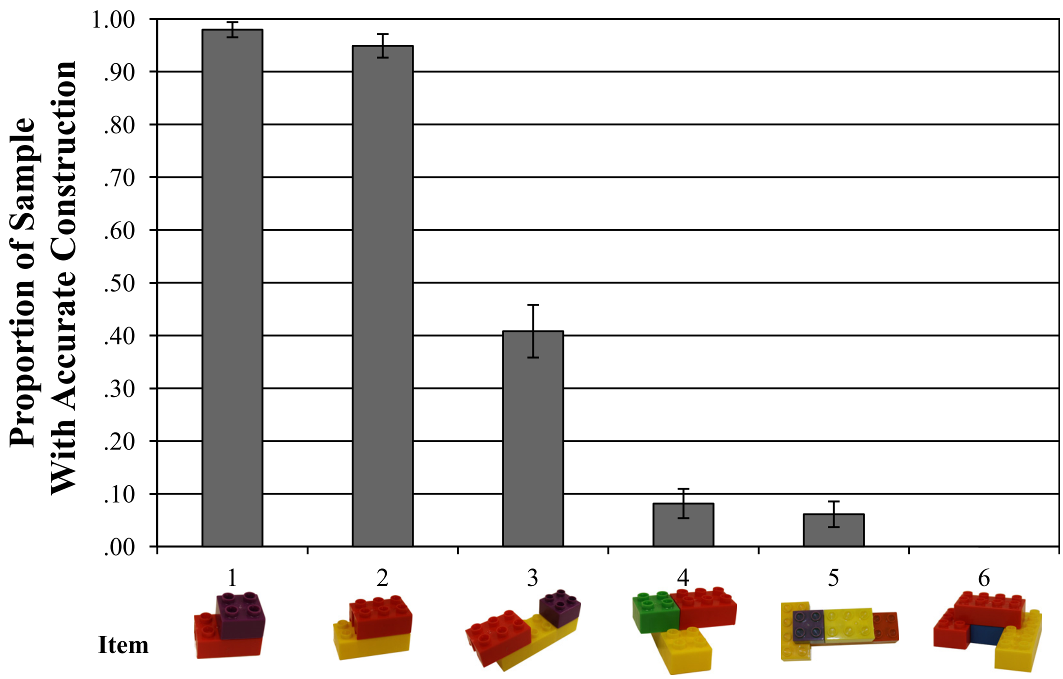


Figure 2. Match scores answer the question: "Did the child's copy match the model completely?" Figure shows Match score by each block design.

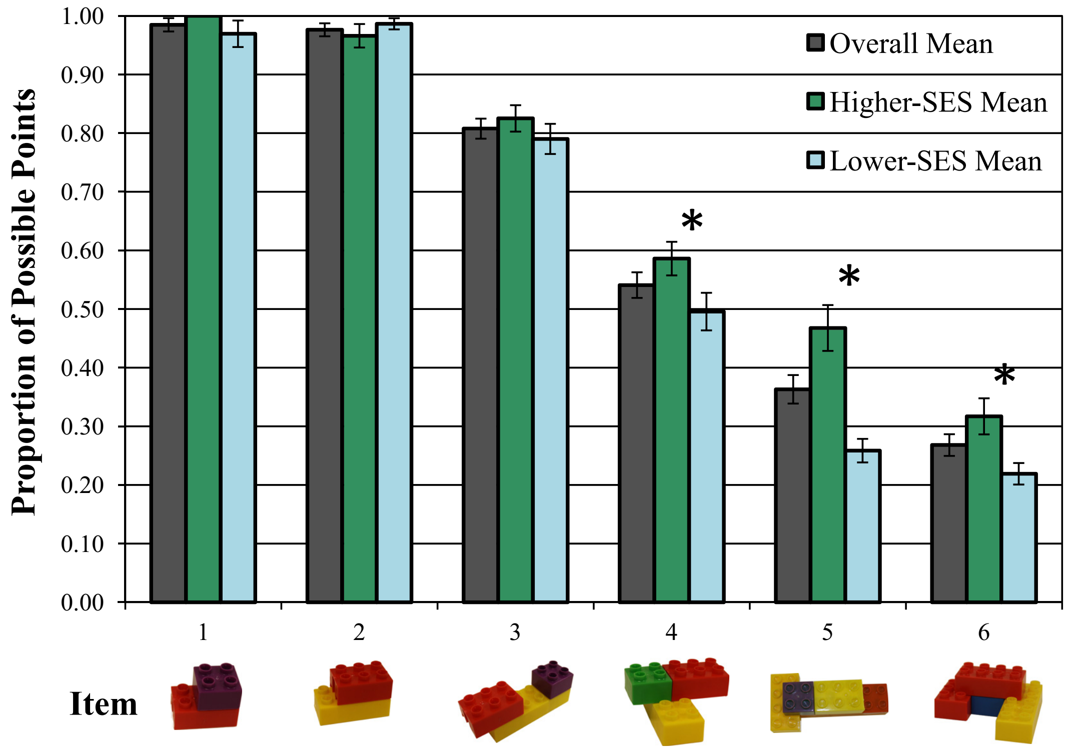


Figure 3. Average of more detailed dimensional scores by block design and socioeconomic status (SES). *SES means different at $p < .05$ level.

the seven children who made errors on Design 1 or 2, none made errors on both, indicating that children understood the task and had the fine motor control needed to succeed. Although these items are presented in graphs, they were dropped from further analyses due to a ceiling effect.

Dimensional scores. Overall, there were similar patterns of performance and correlations for Match scores and the Dimensional scores ($r = .64$, $p < .001$; Figure 3 and Table 2). A 3 (scoring dimension) × 4 (Designs 3–6) repeated measures analysis of variance with gender and SES as between-subjects factors revealed the expected between-subjects effect of SES, $F(1, 94) = 21.4$, $p < .001$, $\eta_p^2 = .19$, with a Bonferroni-corrected post hoc test showing that the higher SES group outperformed the lower SES group (*M*s = .57 and .45, respectively). A Design × SES interaction, $F(3, 252) = 4.2$, $p = .006$, $\eta_p^2 = .043$, indicated that Item 5 showed the largest discrepancy by SES (see Figure 3). We would have expected Item 6 to have the largest difference since it was the hardest; however, floor effects likely attenuated that pattern.

There was a within-subjects main effect of design, $F(3, 252) = 188.7$, $p < .001$, $\eta_p^2 = .67$, and post hoc tests determined that Items 3–6 were all significantly different from one another (see overall means in Figure 3). Considering the observed pattern of responses from the Match scoring scheme, this pattern was expected. There was also a main effect of scoring dimension, $F(2, 188) = 290.2$, $p < .001$, $\eta_p^2 = .76$, such that all scoring dimensions were significantly different from one another (vertical location M = 68% of the possible points, rotation M = 59%, translation M = 27%). Furthermore, there was a Design × Dimension interaction present, $F(6, 564) = 38.1$, $p < .001$, $\eta_p^2 = .29$. The scores for each dimension on each item are presented in Figure 4, and this interaction demonstrates what is apparent from that graph: Certain dimensions were particularly difficult on certain items (e.g., *rotation* on Item 4) and the configuration of the blocks influences the relative difficulty of the dimensions for that item. Figure 3 shows performance on each item by SES group. There were no significant interactions with SES and no gender differences on this task or for any of the other variables.

The internal consistency of the Dimensional scoring on the 3-D and four items was moderately high ($\alpha = .70$), indicating that the items and dimensions are likely measuring similar constructs but are not completely redundant. Raw scores on

Table 2
Correlations and Partial Correlations Controlling for PPVT Scores Between the Design Assembly Task and the EMAS

			Individual dimensions		
EMAS items	TOSA Match score	TOSA Dimensional score	Vertical location	Rotation	Translation
Free count					
r	.49**	.44**	.35**	.37**	.43**
pr	.45**	.39**	.29**	.33**	.38**
Give-N					
r	.28**	.38**	.38**	.24*	.32**
pr	.11	.20	.20	.10	.15
Number order					
r	.37**	.27**	.16	.17	.35**
pr	.30**	.16	.04	.09	.26*
Mental computation					
r	.31**	.33**	.25*	.24*	.31**
pr	.24*	.23*	.15	.17	.23*
EMAS total[a]					
r	.40**	.42**	.35**	.28**	.41**
pr	.30**	.27**	.18	.16	.29**

Note. [a]Combined score of the Give-N, number order, and mental computation subscales. PPVT = Peabody Picture Vocabulary Test; EMAS = Early Mathematics Assessment System; *r* = Pearson correlation; *pr* = partial correlation controlling for PPVT scores; TOSA = Test of Spatial Assembly.
**p* < .05. ***p* < .01.

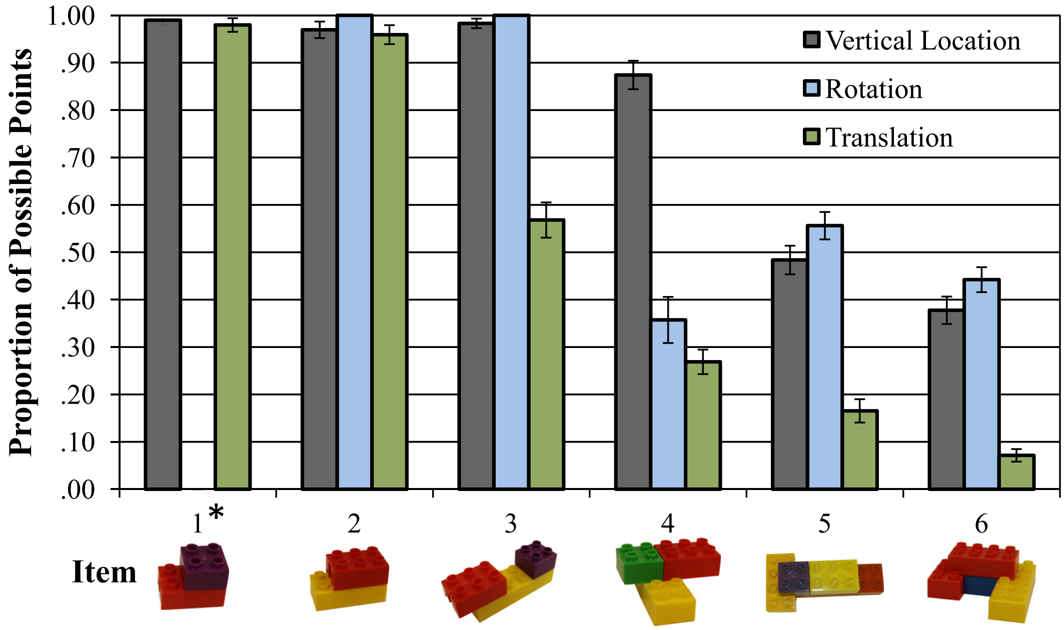


Figure 4. Scores on each block design by dimension. *Rotation was not scored for Item 1 because the purple 2 × 2 component piece attached to on top of the red 3 × 2 base is symmetrical.

the PPVT correlated positively with 3-D TOSA scores across the three dimensions ($r = 0.36$, $p < .001$). To control for language skill, the PPVT was partialed out in most correlation analyses that follow (noted with pr_{ppvt}).

Relation Between Spatial Assembly and Early Mathematics

As expected, even after controlling for the PPVT, Match and Dimensional TOSA scores were correlated with total scores on the EMAS (pr_{ppvt} = .30, p = .004 and pr_{ppvt} = .26, p = .012, respectively) indicating that spatial skill independently predicts a significant amount of the variability in concurrent mathematics performance. The largest correlation with a subset of the EMAS was for the free count portion (pr_{ppvt} = .38, $p < .001$). In fact, the free count was correlated with scores on every 3-D TOSA dimension (see Table 2) and was most highly correlated with Match TOSA scores (pr_{ppvt} = .45, $p < .001$). The only other portion of the EMAS with significant correlations with 3-D TOSA scores was the Mental Computation component (pr_{ppvt} = .23; p = .025) but it was not significantly correlated with scores on any of the individual dimensions.

A stepwise regression was done to determine the contribution of Match scores and Dimensional scores to the prediction of EMAS total scores. Match scores entered the model first, accounting for 15% of the variability in EMAS scores (R = .40, R^2 = .16, $p < .001$, B = 1.68, SE = .40, standardized B = .40, t = 4.23, $p < .001$) and Dimensional scores did not significantly predict additional variance. In an additional regression using only Dimensional scores, Translation was entered into the model first and accounted for 14.6% of the variability in EMAS scores (R = .39, R^2 = .16, $p < .001$, B = .81, SE = .20, standardized B = .39, t = 4.13, $p < .001$). That level of prediction is similar to the amount of variability accounted for by Match scores. Other dimensions did not account for additional EMAS variance in this regression.

Relation Between Spatial Assembly and SES, Gender, and Spatial Language

As expected, the number of relational words that parents said they used with their children differed by SES, with lower SES children exposed to fewer spatial relation words than higher SES children (see Table 3). This difference appears to be driven mainly by four specific words (*between*, *below*, *above*, and *near*). Overall exposure to relational words was significantly correlated with 3-D TOSA scores (r = .33, p = .002) and EMAS scores (r = .42, $p < .001$) in the full sample. These held even after controlling for language skill via the PPVT (TOSA pr_{ppvt} = .22, p = .043; EMAS pr_{ppvt} = .25, p = .021), so these correlations do not likely result from having more overall exposure to language. The sum of responses on only the four words above correlated as well with 3-D TOSA and EMAS scores as the full list (pr_{ppvt} = .23, p = .031; pr_{ppvt} = .25, p = .018, respectively), suggesting that these concepts are particularly salient to task performance. Because two of these words, *above* and *below*, might be expected to be useful in placing pieces on the correct level of the construction, we analyzed them separately. Indeed, exposure to those words was significantly related to vertical location scores (pr_{ppvt} = .22, p = .035) but not to the others.

Table 3
Significant Effects by SES for Spatial Relation Words

By SES	Lower SES *M* (*SD*)	Higher SES *M* (*SD*)	*F*(1, 91)	*p*
Spatial words[a]				
Between	.51 (.51)	.78 (.42)	7.35	.008
Below	.69 (.47)	.86 (.35)	4.25	.042
Above	.58 (.50)	.86 (.35)	9.57	.003
Near	.60 (.50)	.92 (.27)	16.18	.000
Sum of spatial words (of 13)	10.33 (2.92)	11.92 (1.98)	9.07	.003

Note. The full set of words included: *big, between, below, behind, next to, short, little, on, above, near, under, in, long, in front.* SES = socioeconomic status.
[a]Dichotomous responses (1 = *used*, 0 = *not used*).

Discussion

This study was conducted to assess 3-year-olds' spatial assembly skills and the relation between these skills and mathematical skill. We also engaged in a process analysis of children's solution strategies on the TOSA to understand which dimensions were especially difficult for children and which related relatively more to mathematical skill. SES and gender were also considered. Among the key findings is that, while gender differences were not apparent, Match scores indicated that differences between socioeconomic groups on block construction were already evident by the age of 3. The Match scoring scheme was correlated with the more complex Dimensional scoring and had similar correlations with other measures. However, the Dimensional scoring was not a demonstrably better measure of *overall* performance. The primary advantage of the Dimensional scoring was that it yielded important additional information about dimensions of spatial thinking that children find challenging and related differentially to their mathematical skills.

What Have We Learned About Children's Skill in Copying Block Constructions?

Match scoring shows the items had a clear range of difficulty (Figure 2) and that items with more pieces required more analysis of the spatial relations between them. When the designs contained only two pieces, nearly all participants accurately recreated them, a clear indication that children understood the task. High performance on the first two items shows that children used the intrinsic differences (i.e., color, length) between the blocks to arrange them (Uttal et al., 2013). Importantly too, children's success with those designs indicated that they possessed the motor control to complete the task. Only once the spatial relations became more complex as blocks were added did children begin to lose a significant number of points.

As the number of pieces increased, so did the complexity of the spatial relations and number of possible arrangements. Our designs varied in how many blocks appeared on each level, whether the long axes of attached blocks were parallel or perpendicular, and the use of connections with differing overlap between blocks (e.g., four-pip and two-pip connections that only use half the width of a block, etc.). These design characteristics required careful use of the model as it depicted the static and extrinsic relations between the pieces. For example, Items 3 and 4 had the same number and type of blocks, but performance on Item 4 was significantly lower, particularly for the *rotation* and *translation* dimensions. On Item 4, participants often failed to recognize that the blocks on top could share the two-pip width of the base block—possibly because this requires thinking about fractions or using the pips as units. Instead, children either put the pieces on top of the base but parallel to it (a rotation error) or attached both top pieces by four pips and kept the red piece perpendicular to the base (a translation error).

As these common errors reveal, it was typical for one of the spatial relations between blocks to be conserved even though many children were unable to reproduce all relations simultaneously. Thus, children struggle with creating more than one extrinsic relation in a construction. These results are consistent with studies from other domains that show that children have trouble with greater relational complexity (e.g., Halford, Andrews, Dalton, Boag, & Zielinski, 2002) and have difficulty thinking of a single object in multiple ways simultaneously (e.g., Diamond, Carlson, & Beck, 2010).

Interestingly, a training study by Grissmer et al. (2013) has shown that a 28-week intervention based on spatial activities had medium to large effects on *executive function*, part of which entails readily switching between tasks. Building with blocks in 3-D may be important for children's thinking if it increases the number and complexity of relations preschoolers simultaneously represent.

Dimensional Scoring and the Relations Between TOSA Performance and Mathematics

Overall, participants struggled with the dimension of *translation* as the complexity of designs increased, netting < 30% of the total points on each of Designs 4–6 (see Figure 4). Translation errors, by comparison to other errors, are visually subtle and children must be able to notice a shift of a block by only one pip. Even if they do notice these small misalignments, they may not know how to repair them. Our 3-year-olds responded with translation errors as if they lacked an understanding of units and the significance of counting for recreating these relations. Young children have more difficulty counting units when the problem requires switching between parts and wholes (Sophian & Kailihiwa, 1998), something our task requires on multiple levels. That is, the pips are part of the blocks and the blocks are part of the overall construction. If children view the blocks as a "part" and do not recognize that the blocks are also a "whole" segmented by the pips, then they are unlikely to use a counting strategy.

Experience with blocks may allow children to practice decomposing blocks and structures into components, for example, recognizing that larger blocks and block constructions are made up of smaller units. Dividing a whole into units is not something children naturally do. Shipley and Shepperson (1990) showed that when young children count, they have a bias to process discrete physical entities as wholes whether they are parts or not (e.g., broken forks are counted separately). Understanding units with blocks may prepare children for learning that numbers can also be decomposed into smaller units (Clements & Sarama, 2007; NCTM, 2007). Grasping these concepts is important for understanding numbers and lays the foundation for mathematical reasoning about fractions, area, volume, and measurement (Cross et al., 2009). Our data support the conclusion that mathematics is used in block play and particularly for the *translation* dimension, which requires counting units. *Translation* scores correlate with four of the five EMAS measures in Table 2. The more children know about mathematics, the more they succeed on *translation*. Perhaps children who count with caregivers master applying numbers to discreet entities (e.g., horses in a picture book) and then are more likely to move on to counting subunits of whole entities (e.g., the hooves of horses).

The *rotation* dimension also proved to be particularly difficult, with children netting < 60% of the available points for Items 3–6. At first glance, 60% of the points would appear to be relatively good performance, but note that chance responding for the *rotation* dimension is 50%. Although this task is not a pure measure of *mental* rotation because children could rotate the model, difficulty with this dimension is not surprising since children struggle with all but the simplest mental rotation tasks and often use nonrotational strategies to solve problems (e.g., Estes, 1998; Frick & Newcombe, 2009). Similar to *translation*, a failure to sit a piece in its correct orientation could also be related to children's failure to distinguish units based on the pips. A 2 × 3 block looks only marginally different when rotated incorrectly. Therefore, failure to rotate the blocks into the proper position may be intermediate between *translation* and *vertical location* as it relates to mathematics. *Rotation* correlates with the highest number a child could reach in counting ($r = .33$, $p < .05$), suggesting that mathematical skills—and particularly the understanding of units and measurement—may well be implicated in children's failures.

Vertical location was notably easier than the other two dimensions (2-D), although there was room for improvement and errors increased with the complexity of the designs. Even considering that vertical location performance was relatively good, the number of vertical location errors is surprising given that (a) the 3-D model, rather than a fixed 2-D photograph or drawing, was available; (b) the errors seem so apparent to adult eyes; and (c) there are only two levels in each design. *Vertical location* appears to be less related to mathematical skills than *translation*, probably because *vertical location* does not have a metric component. So why do children lose points for *vertical location*? Research with infants suggests that the relations of *over* and *under* between two visible objects are distinguished early on in static displays (Quinn, 1994) and even in dynamic events (e.g., Roseberry, Göksun, Hirsh-Pasek, & Golinkoff, 2012). Perhaps as the complexity of the designs increase, conserving vertical location is swamped by the number of other relations children must maintain. These errors are reminiscent of the errors children make in writing; they keep the letters of their name together, for example, but may confuse left and right or up and down for individual letters (Cornell, 1985).

Note that all of the designs could have been built accurately by simply focusing on the spatial relation between pairs of individual blocks, which almost all children could do for the easiest constructions. However, even if children *could* have segmented the constructions into dyads, this approach probably required more planning and strategy than 3-year-olds can muster without training. This raises the question of whether children could recognize mistakes in other people's designs. Adding a recognition component could extend the age range of the task, providing interesting information about the earlier development of these skills.

Gender, SES, Spatial Language, and Relations to Block Play

As expected, there was no significant gender difference in 3-D TOSA performance. Despite reported gender differences on mental rotation (Linn & Petersen, 1985; Quinn & Liben, 2008), prior studies of structured block building have not found them (Casey et al., 2008).

Although the 3-D TOSA is a nonverbal spatial task, lower SES children performed significantly worse than higher SES children, mostly on the more difficult constructions (4–6). This suggests that lower SES children may have less experience with block play than they do with puzzles (Levine et al., 2012), making block construction more difficult for them. Blocks may not be a purchasing priority for lower SES families when the marketplace is convincing parents that their children need more expensive electronic toys. The fact that children of lower SES families were already worse at the age of 3 is an unfortunate harbinger given the relation between spatial and mathematical skills. And the fact that these are children attending Head Start, a service designed to mitigate SES differences in development, only increases the concern for those not enrolled.

Parent reports of the spatial language they used with their child indicated significant differences between higher and lower SES participants. Lower SES parents indicated that they used fewer spatial words, particularly words that convey spatial relations between two objects (specifically, *between*, *below*, *above*, and *near*) rather than size (e.g., *big* or *short*). Parental use of words that encode spatial properties and features (e.g., the words *edge*, *line*, and *corner*) from 14 to 46 months predict children's spatial skills at 54 months (Pruden et al., 2011). Pruden et al.'s (2011) correlations were from .29 to .48 or, after controlling for parents' overall language production, from .09 to .27. This is similar to the relation between 3-D TOSA scores and parent-reported exposure to spatial relation words that we observed ($r = .33$, $p = .002$ and $pr_{\text{ppvt}} = .22$, $p = .043$). We also found relations between spatial language and EMAS scores ($r = .42$, $p < .001$; $pr_{\text{ppvt}} = .25$, $p = .021$). These results suggest that parental spatial language input is part of an important foundation for STEM learning.

Sociocultural theory (e.g., Vygotsky, 1978) asserts that cognitive development is a product of children's interactions with more knowledgeable social partners, who implicitly tailor instruction to focus just beyond the child's current skill level. Certainly, making parents aware of how important it is to interact with their children during spatial activities should be an important future goal. Furthermore, toy manufacturers may be willing to tout the educational aspects of blocks, include instructions for specific activities that can be implemented to increase parents' use of spatial language and scaffolding, and provide these toys to lower income populations as part of their social outreach. Research indicates that when parents and children play with blocks together toward a common goal, children hear more spatial language than when the play is an open-ended one (Ferrara et al., 2011). Yet parents' role in promoting spatial language and learning needs further exploration. For example, although our sample captures a significant range of the SES spectrum, our measure is admittedly coarse and cannot tease apart the specific factors impacting children. Studies of naturalistic parent–child interactions could clarify the link between parent language and spatial skill and prove important for creating interventions.

Policy makers have more control over early school experiences, where blocks and spatial language can be inserted into the curricula to reverse or prevent a widening of the achievement gap that already exists. And because of the link between spatial and mathematical skills, we can expect that spatial instruction will have a "two-for-one" effect, yielding benefits in mathematics as well.

Conclusion

As the quotation by Emerson that opened the article suggests, children "amidst their baubles" are indeed learning about their effects on the world. However, performance on the TOSA reveals that 3-year-olds have much to learn. Translation errors were common and likely related to not-yet-mastered foundational mathematical skills like counting and measurement. Rotation performance was similarly low, demonstrating again that mental rotation is difficult for young children. Vertical location, despite being significantly easier for children, became a challenge as construction complexity increased.

This study also establishes an early link between spatial and mathematical skills, with other research suggesting that this link strengthens with time (e.g., Battista, 1990; Wolfgang et al., 2001), perhaps unfortunately for children from lower SES backgrounds who are already falling behind in these skills by age 3. However, blocks are relatively affordable and block-building activities can be structured as a cooperative social activity that facilitates the use of spatial language both between children and between children and their teachers. Thus, there appears to be little practical reason that blocks cannot be used successfully even in poorly funded schools. Additional research may support the view that block building can provide a ripe tool for intervention.

References

Battista, M. T. (1990). Spatial visualization and gender differences in high school geometry. *Journal for Research in Mathematics Education*, *21*, 47–60.

Burchinal, M., Nelson, L., Carlson, M., & Brooks-Gunn, J. (2008). Neighborhood characteristics, and child care type and quality. *Early Education and Development*, *19*, 702–725. doi:10.1080/10409280802375273

Caldera, Y. M., Culp, A. M., O'Brien, M., Truglio, R. T., Alvarez, M., & Huston, A. C. (1999). Children's play preferences, construction play with blocks, and visual-spatial skills: Are they related? *International Journal of Behavioral Development*, *23*, 855–872. doi:10.1080/016502599383577

Casey, B. M., Andrews, N., Schindler, H., Kersh, J. E., Samper, A., & Copley, J. (2008). The development of spatial skills through interventions involving block building activities. *Cognition and Instruction*, *26*, 269–309. doi:10.1080/07370000802177177

Casey, B. M., & Bobb, B. (2003). The power of block building. *Teaching Children Mathematics*, *10*, 98–102.

Chatterjee, A. (2008). The neural organization of spatial thought and language. *Seminars in Speech and Language*, *29*, 226–238. doi:10.1055/s-0028-1082886

Cheng, Y. L., & Mix, K. S. (2012). Spatial training improves children's mathematics ability. *Journal of Cognition and Development*. Advance online publication. doi:10.1080/15248372.2012.725186

Clements, D. H., & Sarama, J. (2007). Effects of a preschool mathematics curriculum: Summative research on the Building Blocks project. *Journal of Research in Mathematics Education*, *38*, 136–163.

Cornell, J. M. (1985). Spontaneous mirror-writing in children. *Canadian Journal of Psychology*, *39*, 174–179. doi:10.1037/h0080122

Cross, C. T., Woods, T. A., & Schweingruber, H. (Eds.). (2009). *Mathematics learning in early childhood: Paths toward excellence and equity*. Washington, DC: National Academies Press.

Diamond, A., Carlson, S. M., & Beck, D. M. (2010). Preschool children's performance in task switching on the dimensional card sort task: Separating the dimensions aids the ability to switch. *Developmental Neuropsychology*, *28*, 689–729. doi:10.1207/s15326942dn2802_7

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test* (4th ed.). Bloomington, MN: Pearson Assessments.

Entwisle, D., & Astone, N. (1994). Some practical guidelines for measuring youth's race/ethnicity and socioeconomic status. *Child Development*, *65*, 1521–1540. doi:10.2307/1131278

Estes, D. (1998). Young children's awareness of their mental activity: The case of mental rotation. *Child Development*, *69*, 1345–1360. doi:10.1111/j.1467-8624.1998.tb06216.x

Ferrara, K., Hirsh-Pasek, K., Newcombe, N. S., Golinkoff, R. M., & Shallcross Lam, W. (2011). Block talk: Spatial language during block play. *Mind, Brain, and Education*, *5*, 143–151. doi:10.1111/j.1751-228X.2011.01122.x

Frick, A., & Newcombe, N. S. (2009). *Measuring mental rotation in 4-year-olds using a nonverbal touch screen paradigm*. Poster presented at the biennial meeting of the Cognitive Development Society, San Antonio, TX.

Fuson, K. C., & Murray, C. (1978). The haptic-visual perception, construction, and drawing of geometric shapes by children aged two to five: A Piagetian extension. In R. Lesh & D. Mierkiewicz (Eds.), *Concerning the development of spatial and geometric concepts* (pp. 49–83). Columbus, OH: ERIC.

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition*, *44*, 43–74. doi:10.1016/0010-0277(92)90050-R

Ginsburg, H. P., Lee, Y.-S., Pappas, S., Hartman, G., & Rosenfeld, D. (2010, June). A comprehensive mathematics assessment for preschool-age children. Poster presented at the biennial meeting of the Head Start's National Research Conference, Washington, DC.

Ginsburg, H. P., Pappas, S., & Lee, Y.-S. (2012). *Early mathematics assessment system (EMAS)*. In preparation.

Göbel, S., Walsh, V., & Rushworth, M. F. S. (2001). The mental number line and the human angular gyrus. *NeuroImage*, *14*, 1278–1289. doi:10.1006/nimg.2001.0927

Grissmer, D. W., Mashburn, A. J., Cottone, E., Chen, W. B., Brock, L. L., Murrah, W. M., et al. (2013, April). *Play-based after-school curriculum improves measures of executive function, visuospatial and math skills and classroom behavior for high risk K-1 children*. Paper presented at the Society for Research in Child Development, Seattle, WA.

Gunderson, E. A., Ramirez, G., Beilock, S. L., & Levine, S. C. (2012). The relation between spatial skill and early number knowledge: The role of the linear number line. *Developmental Psychology*, *48*, 1229–1241. doi:10.1037/a0027433

Halford, G. S., Andrews, G., Dalton, C., Boag, C., & Zielinski, T. (2002). Young children's performance on the Balance Scale: The influence of relational complexity. *Journal of Experimental Child Psychology*, *81*, 417–445. doi:10.1006/jecp.2002.2665

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Brookes.

Hart, B., & Risley, T. R. (2003). The early catastrophe: The 30 million word gap by age 3. *American Educator*, *27*, 4–9.

Hirsh-Pasek, K., Kochanoff, A., Newcombe, N. S., & de Villiers, J. (2005). Using scientific knowledge to inform preschool assessment: Making the case for "empirical validity." *SRCD Social Policy Report*, *19*, 1–19.

Hoff, E. (2013). Interpreting the early language trajectories of children from low-SES and language minority homes: Implications for closing achievement gaps. *Developmental Psychology*, *49*, 4–14. doi:10.1037/a0027238

Jordan, N. C., Kaplan, D., Nabors Oláh, L., & Locuniak, M. N. (2006). Number sense growth in kindergarten: A longitudinal investigation of children at risk for mathematics difficulties. *Child Development*, *77*, 153–175. doi:10.1111/j.1467-8624.2006.00862.x

Kamii, C., Miyakawa, Y., & Kato, Y. (2004). The development of logico-mathematical knowledge in a block-building activity at ages 1-4. *Journal of Research in Childhood Education*, *19*, 44–57. doi:10.1080/02568540409595053

Karp, S. A., & Konstadt, N. (1971). *Children's Embedded Figures Test*. Palo Alto, CA: Consulting Psychologists Press.

Levine, S. C., Jordan, N. C., & Huttenlocher, J. (1992). Development of calculation abilities in young children. *Journal of Experimental Child Psychology*, *53*, 72–103. doi:10.1016/S0022-0965(05)80005-0

Levine, S. C., Ratliff, K. R., Huttenlocher, J., & Cannon, J. (2012). Early puzzle play: A predictor of preschoolers' spatial transformation skill. *Developmental Psychology*, *48*, 530–542. doi:10.1037/a0025913

Levine, S. C., Vasilyeva, M., Lourenco, S. F., Newcombe, N. S., & Huttenlocher, J. (2005). Socioeconomic status modifies the sex difference in spatial skill. *Psychological Science*, *16*, 841–845. doi:10.1111/j.1467-9280.2005.01623.x

Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development*, *56*, 1479–1498. doi:10.2307/1130467

Mix, K. S., Moore, J. A., & Holcomb, E. (2011). One-to-one play promotes numerical equivalence concepts. *Journal of Cognition and Development*, *12*, 463–480. doi:10.1080/15248372.2011.554928

Moore, D. S., & Johnson, S. P. (2008). Mental rotation in human infants: A sex difference. *Psychological Science*, *19*, 1063–1066. doi:10.1111/j.1467-9280.2008.02200.x

National Council of Teachers of Mathematics. (2007). *Second handbook of research on mathematics teaching and learning*. Washington, DC: Author.

Newcombe, N. S. (2010). Picture this: Increasing math and science learning by improving spatial thinking. *American Educator*, *34*, 29–35.

Piaget, J., & Inhelder, B. (1969). *The early growth of logic in the child*. New York: Norton.

Pruden, S. M., Levine, S. C., & Huttenlocher, J. (2011). Children's spatial thinking: Does talk about the spatial world matter? *Developmental Science*, *14*, 1417–1430. doi:10.1111/j.1467-7687.2011.01088.x

Quinn, P. C. (1994). The categorization of above and below spatial relations by young infants. *Child Development*, *65*, 58–69. doi:10.1111/j.1467-8624.1994.tb00734.x

Quinn, P. C., & Liben, L. S. (2008). A sex difference in mental rotation in young infants. *Psychological Science*, *19*, 1067–1070. doi:10.1111/j.1467-9280.2008.02201.x

Robinson, N. M., Abbott, R. D., Berninger, V. W., & Busse, J. (1996). The structure of abilities in math-precocious young children: Gender similarities and differences. *Journal of Educational Psychology*, *88*, 341–352. doi:10.1037//0022-0663.88.2.341

Roseberry, S., Göksun, T., Hirsh-Pasek, K., & Golinkoff, R. M. (2012). Carving categories in a continuous world: Preverbal infants discriminate categorical changes before distance changes in dynamic events. *Spatial Cognition and Computation*, *12*, 231–251.

Saxe, G. B. (1979). Developmental relations between notational counting and number conservation. *Developmental Psychology*, *50*, 180–187.

Shipley, E. F., & Shepperson, B. (1990). Countable entities: Developmental changes. *Cognition*, *34*, 109–136. doi:10.1016/0010-0277(90)90041-H

Sophian, C., & Kailihiwa, C. (1998). Units of counting: Developmental changes. *Cognitive Development*, *13*, 561–585. doi:10.1016/S0885-2014(98)90007-7

Stiles, J., & Stern, C. (2009). Developmental change in spatial cognitive processing: Complexity effects and block construction performance in preschool children. *Journal of Cognition and Development*, *2*, 157–187. doi:10.1207/S15327647JCD0202_3

Stiles-Davis, J. (1988). Developmental change in young children's spatial grouping activity. *Developmental Psychology*, *24*, 522–531. doi:10.1037/0012-1649.24.4.522

Umiltà, C., Priftis, K., & Zorzi, M. (2009). The spatial representation of numbers: Evidence from neglect and pseudoneglect. *Experimental Brain Research*, *192*, 561–569. doi:10.1007/s00221-008-1623-2

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., et al. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin*, *139*, 352–402. doi:10.1037/a0028446

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin*, *117*, 250–270. doi:10.1037/0033-2909.117.2.250

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Wolfgang, C. H., Stannard, L. L., & Jones, I. (2001). Block play performance among preschoolers as a predictor of later school achievement in mathematics. *Journal of Research in Childhood Education*, *15*, 173–180. doi:10.1080/02568540109594958

Woodcock, R. W., Mather, N., & McGrew, K. S. (2001). *Woodcock–Johnson III Tests of Cognitive Abilities*. Itasca, IL: Riverside.

Wynn, K. (1990). Children's understanding of counting. *Cognition*, *36*, 155–193. doi:10.1016/0010-0277(90)90003-3